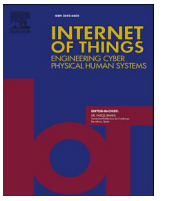
Internet of Things 27 (2024) 101291

Contents lists available at ScienceDirect

Internet of Things

journal homepage: www.sciencedirect.com/journal/internet-of-things



Review article

# The EU's AI act: A framework for collaborative governance

Celso Cancela-Outeda [1]

*Celso Cancela-Outeda, Universidade de Vigo, Spain*

ARTICLE INFO

*Keywords:*
Popularization
Civil society
Stakeholders
Governance
"black box"
Collaborative logic

ABSTRACT

In February 2024, the Council and the European Parliament (EP) agreed on the Artificial Intelligence Regulation (usually known as AI Act, AIA) .[2] This regulation evaluates AI applications to ensure they are used ethically and responsibly, promoting the development of safe and lawful AI across the EU's single market. It establishes a comprehensive legal framework with a risk-based approach, aiming to achieve a balance between protecting the health, safety, and fundamental rights of European citizens and ensuring that the growing AI industry in Europe remains competitive and continues to innovate. The AIA also includes governance mechanisms oriented towards achieving effective implementation throughout the EU. For this purpose, a European Artificial Intelligence Office has already been established. In accordance with the provisions of the forthcoming AIA, it will establish a European Artificial Intelligence Board, an advisory forum, and a scientific panel. Furthermore, it will be set up at the national level the so-called national competent authorities. In this way, a single European governance system for AI is emerging, inspired by collaborative governance, which is essential for achieving fair and effective implementation of AI regulations across the EU. The main objective of this text is to critically examine the governance system established by the AIA. Using the contents of the current version of the AIA (April 2024), this analysis delves into the mechanisms and structures designed to implement AI across the EU. As a conclusion, it offers a critical perspective on the collaborative governance, highlighting its strengths and potential areas for improvement.

## 1. Introduction

Since at least 2018, the field of AI has been experiencing explosive growth. In less than a decade, AI has undergone rapid advancements and has emerged as a critical force shaping our societies at an accelerating pace. Consequently, AI is considered an exponential, disruptive, and cross-cutting technology that raises ethical, legal, and political questions and challenges. These challenges are heavily influenced by one's perspective on AI, whether techno-optimistic or techno-pessimistic [9,10]. This acceleration culminated in a surge of public interest by the end of 2022, fuelled by the emergence of powerful AI models like ChatGPT and DALL-E. The proliferation of numerous AI tools, available in free and commercial versions and applicable across different public and private spheres, has sparked debates on its evolution, impact, use, and regulation.

*E-mail address:* ccancela@uvigo.gal.

[1] Holder Jean Monnet Chair European Institutions and Citizens: Adapting the EU to the Digitalisation Challenges (EICAEUD).

[2] We use the AI Act version P9_TA(2024)0138 Artificial Intelligence Act European Parliament legislative resolution of 13 March 2024 on the proposal for a regulation of the European Parliament and of the Council on laying down harmonised rules on Artificial Intelligence (Artificial Intelligence Act) and amending certain Union Legislative Acts (COM(2021)0206–C9-0146/2021–2021/0106(COD)). https://www.europarl.europa.eu/doceo/document/TA-9-2024-0138_EN.pdf

https://doi.org/10.1016/j.iot.2024.101291

Available online 20 July 2024

Experts view AI as an unstoppable force that is reshaping the trajectory of human history. In the digital age, AI is driving rapid and consequential transformations across numerous social spheres—including politics, economics, law, ethics, and education—with profound implications for the future [27]. In fact, large global companies are at the forefront of AI research, development, and deployment.[3] Governments, on the other hand, are exploring how AI can transform crucial areas like healthcare, economy, or security. Potential benefits include increased efficiency in public administration, optimized service delivery, and solutions to challenges like climate change and public safety. However, while acknowledging AI's transformative potential, governments also recognize its associated risks, which encompass negative impacts on employment, ethical values, privacy, intellectual property, and equality—fundamental rights in general. These concerns have prompted governments to explore diverse regulatory approaches, both direct and indirect, to oversee and regulate AI development.

For some time, regulatory approaches focused on an ethical dimension, advocating for self-regulation or responsible use of AI by companies and users [34]. Later, the focus has been shifting towards legal regulation (binding legislation) (Birkstadt *et al.* [8]. This aims to not only protect against potential misuse of AI but also foster an ethical and sustainable environment for innovation. Ultimately, legal frameworks will provide certainty, incentivize investment and innovation, while mitigating potential risks [29]. The EU exemplifies this approach with its forthcoming AI regulation, which incorporates a dedicated chapter outlining a collaborative governance structure that leverages both European and national institutions while fostering stakeholder participation. Furthermore, building upon this collaborative approach, the regulation is inspired by a human-centred vision. As articulated by Carriço [11], this vision underscores the importance of "progress should always be driven with a human-centred perspective in mind, i.e. one that seeks to augment human intelligence and capability, not replace them.".

The European AIA breaks new ground with a dedicated chapter on collaborative governance that leverages both European and national institutions, while actively encouraging stakeholder participation. The approach draws on a collaborative governance notion,[4] characterized by an ongoing exchange of knowledge between public institutions or agencies and diverse stakeholders (citizens, businesses, NGOs, experts, etc.) in policymaking [7]. This scheme governance transcends mere consultation, fostering a collaborative decision-making process. This approach enables several advantages: it promotes dialogue and integrates diverse viewpoints, leading to the development of more informed policies. Furthermore, collaborative governance fosters a more inclusive and collaborative environment that fosters knowledge sharing, ultimately increasing trust and legitimacy. However, implementing collaborative governance effectively presents its own set of challenges. Ensuring equitable representation is a key goal, since influential stakeholders can otherwise dominate the process (e.g. "capture" policymaking).

This article examines the landscape of AI regulation in the EU, with a particular emphasis on the collaborative governance mechanisms. Embracing this approach for AI is not just relevant, but necessary for several reasons. Firstly, the huge complexity of AI demands a multi-stakeholder approach, drawing on expertise from diverse fields. for crafting effective and comprehensive AI regulations. So, collaboration ensures a holistic understanding. Secondly, the fast pace of AI development demands a flexible and adaptable regulatory framework. A collaborative perspective allows for continuous stakeholder engagement, fosters ongoing dialogue, and ensures that regulations can be adjusted as AI technology evolves.

This contribution dives into the necessity for EU action on AI, followed by a historical analysis of the key steps leading to the current AIA. At the core lies the recently established European AI Office (launched in early 2024), envisioned as the cornerstone of a single European approach to AI governance. We will then explore its institutional function and its collaborative work with other key players, including the European Artificial Intelligence Board, the Commission's European Centre for Algorithmic Transparency (ECAT), and national authorities.

## 2. AI legislation in context: understanding the broad implications

AI has shed its futuristic skin and become a tangible force reshaping society across various sectors. A landmark development in this journey was the launch of ChatGPT, which sparked a rapid emergence of diverse and powerful AI tools. These tools excel at a wide range of tasks, and at their core lies a transformative power: the ability to significantly improve efficiency and productivity, ultimately impacting economic growth. While AI has demonstrably impacted various sectors, its most significant effects, and their magnitude, remain uncertain. To assess the potential for AI-driven productivity improvements, a comparison can be drawn with historical examples of general-purpose technologies. These technologies, exemplified by the steam engine and electricity, share key characteristics: widespread application across the economy, sustained performance improvements within the technology itself, and the ability to spur innovation in industries that adopt them, ultimately leading to broad-based productivity growth. AI's versatility and adaptability across applications suggest its potential as a general-purpose technology, as evidenced by expectations of broad productivity improvements in the future [5]. Furthermore, generative AI applications, such as creating marketing copy or generating realistic images, are already being marketed to become a hugely lucrative business [4].

As Singh [31] points out "the technological leaps of the last decade that have ushered in the era of AI mass production". AI has demonstrated its potential to foster development in a variety of fields. From automating everyday tasks to solving problems in advanced scientific fields, its impact is vast and diverse [2]. Beyond automating tasks, AI is transforming business by enabling

[3] Among the leading AI companies in the world are: OpenAI, Google, IBM, Microsoft, NVIDIA, Amazon, Anthropic, Anduril, DeepMind, Baidu, Tencent, Huawei, etc.

[4] Bianchi and al. (2021) writes: "Such phenomenon also is witnessed by the variegated terminology contained in the literature around related concepts encompassing public governance, networks, collaboration, and public value".

applications like AI-powered chatbots for 24/7 customer support and market analysis tools for data-driven decision-making. AI's influence in the industrial sector is multifaceted. It drives advanced robotics, enabling more accurate and efficient operations. Furthermore, AI-powered predictive maintenance helps anticipate equipment failures, minimizing downtime and maximizing production. AI can analyse processes, optimizing them for increased efficiency and productivity, too.

AI's transformative influence extends beyond automation. In healthcare, it revolutionizes diagnostics by improving accuracy and speeds up drug development [20,26]. Moreover, AI facilitates personalized medicine (precision medicine) through chatbots, making healthcare more accessible [1]. The financial sector benefits from AI's ability to detect fraud, protecting both institutions and customers [6]. In addition, AI-powered financial advisors provide personalized investment recommendations [6]. High-frequency trading leverages AI algorithms for rapid transactions, capitalizing on market inefficiencies. In the transportation and logistics sector, AI plays a dual role. It contributes to the development of safer and more efficient autonomous vehicles [25], while also optimizing traffic management in cities, reducing congestion and improving overall flow.

In agriculture, AI facilitates precision farming by optimising the use of resources (water, fertilisers, pesticides, etc.), which increases productivity and reduces environmental impact. Precisely in this field, AI is used to analyse data from satellites and other sources to monitor climate change and its impacts, improving scientific understanding and developing appropriate strategies. It is also used in the management of natural resources (water, forests, fisheries, etc.), optimising their use (precision agriculture) and ensuring their sustainability through the analysis of available data [16,19].

AI also presents itself as a valuable tool in education, offering personalized tutoring and automated assessments, freeing teachers' time for more impactful student interaction and feedback [3,13]. However, it's important to ensure a balanced approach. In this field, over-reliance on AI in the classroom can discourage personal interaction between students and teachers and limit the development of social and critical skills. AI could lead to the limitation of students' creativity and critical thinking if it is used to automate tasks that they used to do themselves. It could dehumanise the learning process and the relationship between students and teachers, reducing or eliminating human interaction and empathy [30].

AI's influence extends to the entertainment and information sectors. Generative AI personalizes content discovery by recommending material based on user preferences, simplifying the search process. However, its ability to create content raises concerns about "deepfakes" and the potential for manipulating public opinion. Furthermore, generative AI's capabilities extend beyond content discovery. It can create human-quality text, images, and even audio, blurring the lines between human and machine-generated content. This innovation has wide-ranging implications. On the one hand, it represents a revolutionary step in contents creation. On the other hand, it raises ethical concerns regarding intellectual property rights, information accuracy, and the representation of people [14]. Moreover, the opacity of AI-generated content creates challenges in determining accountability and transparency. Who is responsible for the content generated by AI? These issues highlight the potential impact of AI on areas like privacy, discrimination, and even the functioning of democratic systems due to the power it grants technology companies.

AI's influence extends to the legal field, assisting with tasks like predicting judicial outcomes by analysing past judgments and identifying relevant factors [12]. AI can also support crime detection by analysing large datasets, such as facial recognition in security footage, to identify patterns that may be indicative of criminal activity. These applications showcase the vast potential of AI across various domains, including quantum computing, astronomy, physics, and chemistry (all easily scalable). As AI technology advances, we can expect even more innovative applications with far-reaching economic, legal, social, and societal implications [32]. This progress will undoubtedly prompt new, and potentially controversial, questions.

The preceding text reflects an optimistic view of AI applications across various fields, primarily highlighting its benefits in terms of efficiency and productivity. However, it scarcely mentions the potential risks and problems associated with the use of AI (job losses, greater socio-economic inequality, labour retraining, wealth redistribution, algorithmic bias, discriminatory and unfair decisions, errors or failures, lack of transparency, and so on.). In the end, such aspects can undermine public trust in these technologies [32]. In this way, unveiling the multifaceted nature of AI, we discover a series of implications that point towards a potential dark side [24]. To address the potential threats, we can categorize the main sources of catastrophic AI risks into four key groups: 1) malicious use, where individuals or groups deliberately employ AI to cause harm; 2) AI race, characterized by competitive environments that drive actors to deploy unsafe AI systems or relinquish control to AI; 3) organizational risks, which underscore how human factors and complex systems can heighten the likelihood of catastrophic accidents; and 4) rogue AIs, describing the inherent challenge in controlling agents that possess intelligence surpassing that of humans [18].

The recent exponential growth in AI development and commercial applications across various sectors [4,15] is fuelled by several interrelated factors. Firstly, advances in hardware enable efficient processing of vast datasets, paving the way for more sophisticated algorithms. Secondly, global connectivity and information digitization provide AI with a rich foundation for learning and improvement (large datasets). Thirdly, collaborations between academia, industry, and research institutions foster innovation and knowledge sharing. Fourthly, increased investments from technology and financial firms accelerate research and development in AI, further driving commercial applications and investments. Finally, the growing demand for AI across industries (healthcare, logistics, etc.) creates a continuously expanding market. Public awareness of AI's potential and societal impact further fuels its rapid growth, with increased support from governments and businesses.

In March 2023, hundreds of leading AI experts, technology entrepreneurs and scientists signed an open letter calling for a pause in the development (suspended training) and testing of AI technologies more powerful than OpenAI's GPT-4 language model.[5] They hope

[5] See, Pause Giant AI Experiments: An Open Letter. https://futureoflife.org/open-letter/pause-giant-ai-experiments/ (last access 23 April 2024).

to have the time to adequately study the ethical and security risks. They also warned of the dangers of the "race out of control" to develop ever bolder AI systems. The pace of these transformations goes hand in hand with huge investments in AI. Large tech companies, as well as innovative startups, are investing huge sums of money in research and development to drive the advancement of AI. This increased investment is fuelling fierce competition for leadership in the field, which in turn is driving an accelerated pace of innovation. The rapid development and innovation of AI far outpaces the speed at which standards are adopted, posing a challenge when implementing regulations.

While the fast pace of AI development presents a challenge, its inherent complexity creates a formidable obstacle for policymakers seeking to regulate its use. This "black box" nature, where internal decision-making processes are opaque, also poses significant difficulties for effective regulation. Policymakers not only struggle to understand how AI reaches conclusions but also pinpoint accountability for AI-driven outcomes and adapt regulations to keep pace with AI's rapid evolution. In this complex landscape, collaboration with experts is a cornerstone. Acknowledging the "black box" nature of AI is essential for policymakers to develop a nuanced regulatory approach, ensuring AI deployment is effective.

Another note regarding AI is its popularization, which can be described as dizzying [28]. Although AI was already a reality in the hands of certain professionals, the introduction of ChatGPT at the end of 2022 caused a tremendous increase in the number of users in a matter of months. As of early 2023, some emerging generative AI systems had reached more than 100 million users [15]. But, as Franganillo (2024) points out, "although AI is quickly reaching all audiences, it cannot be said that it is being democratized, since these audiences must not only be able to use it, but they must also be able to understand it. […] Society in general still finds it difficult to understand what AI is and what it is really capable of".

To sum up, the accelerated growth of AI, characterized by efficiency gains, exponential expansion, and widespread adoption, needs the development of legal frameworks to address its inherent risks and ethical implications. But, the inherent complexity of AI systems, often described as "black box" poses a significant challenge to effective regulation. Experts warn of challenges related to ethics, privacy, and personal safety as AI becomes increasingly ubiquitous. Specific AI legislation is imperative to establish a regulatory framework governing the rights and responsibilities of developers. This framework should be balanced, ensuring the protection of individual rights, the establishment of ethical limitations, and the promotion of algorithmic transparency [33]. In conclusion, as AI applications proliferate across various public and private sectors, it has transitioned from a technological innovation to a subject demanding effective governance [8].

## 3. The EU's path towards effective regulation AI

In a few weeks, the EU will have the first specific and comprehensive regulation of AI. The Council and the EP reached a provisional agreement on the legal text on 9 December 2023, after months of negotiations (trilogue). The EU's AIA represents a groundbreaking piece of legislation, being the world's first of its kind. This landmark regulation aims to harmonise rules governing AI systems to ensure their safety and compliance with fundamental rights and EU values. However, this achievement marks the culmination of various initiatives adopted gradually at the European level, particularly led by the European Commission. Let us briefly review the main initiatives.

Although lagging behind the US and China, the EC officially recognised the strategic importance of AI and data from both an economic and geostrategic perspective and published in February 2020, the White Paper on Artificial Intelligence and the European Data Strategy.[6] Earlier, in 2018,[7] the EC had set up the High-Level Expert Group on AI (IA HLEG)[8] to discuss ethical, legal, economic and social aspects. It gathered expert input and rallied a broad alliance of diverse stakeholders and drafted Ethics Guidelines for AI.[9] It was tasked with advising the EC on building a broad and diverse community of stakeholders in a "European AI Alliance",[10] supporting the implementation of the upcoming European initiative on AI, and drafting guidelines for the ethical development and use of AI based on the EU's fundamental rights. It aimed to consider issues such as fairness, safety, transparency, democracy, and more broadly, the impact on the application of the Charter of Fundamental Rights. Some months later, the Expert Group presented the Policy and Investment Recommendations for Trustworthy AI[11] that provides further guidance and recommendations on this issue. Also in February 2018, EU members states (and Norway) signed a Declaration of cooperation on AI,[12] expressing a strong will to join forces for a European approach. They aim to ensure Europe's competitiveness in AI research and deployment, while collectively addressing social, economic, ethical, and legal questions. Following this declaration, AI ascended to the forefront of European policy agenda, bolstered by significant political support.

[6] White Paper On Artificial Intelligence - A European approach to excellence and trust COM(2020) 65 final https://op.europa.eu/en/publication-detail/-/publication/ac957f13-53c6-11ea-aece-01aa75ed71a1 (last access 12-05-24).

[7] In October 2017, the European Council asked the European Commission to present a European approach to AI orientated to a Digital Europe. This European institution called for a 'sense of urgency to address emerging trends' including 'issues such as artificial intelligence…, while at the same time ensuring a high level of data protection, digital rights and ethical standards'. European Council meeting (19 October 2017) Conclusions. https://www.consilium.europa.eu/media/21620/19-euco-final-conclusions-en.pdf

[8] *High-Level Expert Group on Artificial Intelligence* https://ec.europa.eu/digital-single-market/en/high-level-expert-group-artificial-intelligence

[9] Ethics Guidelines for Trustworthy Artificial Intelligence, https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=60419

[10] European AI Alliance, https://digital-strategy.ec.europa.eu/en/policies/european-ai-alliance

[11] See, https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=60343 (last access 20-04-24)

[12] Declaration of Cooperation on AI. https://ec.europa.eu/newsroom/dae/redirection/document/50951

In April 2018, the Juncker Commission published the European AI Strategy[13] setting out Europe's own path for future regulation, while characterising AI as one of the "most strategic technologies of the 21st century". Based on this document, in December 2018, the EC published the Coordinated Plan on AI[14] which aimed at joint actions for closer cooperation between EU member states (plus Norway and Switzerland) and the European Commission. This document defined four key areas: increasing investment, increasing data availability, promoting talent, and developing ethical and trustworthy AI. This Plan was reviewed in April 2021. It aims to solidify EU leadership in trustworthy AI and focuses on accelerating AI investments for economic and social recovery, implementing AI strategies promptly, and aligning AI policy to address global challenges and remove fragmentation.[15]

In 2019, the AI HLEG published a set of ethical guidelines. This document aimed to inform future regulatory actions and promote the development of trustworthy AI.[16] According to these Guidelines, trustworthy AI should be lawful (respecting all applicable laws and regulations), ethical (adhering to ethical principles and values), and robust (both technically and in consideration of its social environment). In the same year, the EC published the Communication: Building trust and confidence in people-centred artificial intelligence.[17] This document mentions some key requirements for trustworthy AI, including ensuring human agency and oversight, technical robustness and safety, privacy and data governance, transparency, and accountability. These elements are essential for building AI systems that users can rely on and trust. The Guidelines were complemented by the Assessment List for Trustworthy AI, which serves as the operational tool. It operationalizes the key requirements for ethical AI and provides practical guidance (a checklist) for implementation by developers and deployers. On July 2020, the AI HLEG presented their final Assessment List for Trustworthy AI (ALTAI) .[18]

In February 2020, the EC published the White Paper on Artificial Intelligence: a European approach to excellence and trust.[19] On the one hand, it assumed that the promotion of digitalisation is a cardinal factor in fostering competitiveness and boosting European progress and social welfare; on the other hand, it also acknowledged the economic and geostrategic significance of AI and data. For this reason, this document was accompanied by the European Data Strategy. Both documents were the first pillars of the Digital strategy presented in the European Data Strategy[20] with the intention of advancing the development of a common data space where data can circulate in all sectors; in particular, the so-called "common European data spaces in strategic sectors and domains of public interest" (industrial, European Green Deal, mobility, health, financial, energy, agriculture and data for public administration). In addition to contributing to European integration, it is crucial for the development of AI in the EU, as data is, after all, the raw material that feeds AI.

On 23 July 2020, the EC presented a Proposal for a legal act of the EP and the Council laying down requirements for AI.[21] In April 2021, in order to complement the mentioned Coordinated Plan, the EC proposed the first-ever legal framework on AI. This Proposal for a Regulation on AI categorizes AI uses into four levels of risk: unacceptable, high, limited, and minimal. The regulation aims to ensure that Europeans can trust the AI they use, while also fostering an ecosystem of excellence in AI and enhancing the EU's global competitiveness.[22] At the same time, the EC presented the Communication on Fostering a European approach to AI, which outlines the EU's strategy for AI, emphasizing trustworthiness, competitiveness, and ethical considerations. It aims to ensure that Europe remains at the forefront of AI development while safeguarding citizens' rights and values.[23]

In 2021, the EC launched the public consultation on Civil liability, which seeks to adapt liability rules to the digital age and AI. It

[13] Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions Artificial Intelligence for Europe [COM(2018) 237 final]. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=COM:2018:237:FIN (last access 27-04-2024).

[14] Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions Coordinated Plan on Artificial Intelligence, COM/2018/795 final. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52018DC0795 (last access 15-04-2024).

[15] The 2021 Coordinated Plan on Artificial Intelligence. https://ec.europa.eu/newsroom/dae/redirection/document/75787 (last access 23-04-2024)

[16] Ethics guidelines for trustworthy AI. https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=60419 (last access 27-04-24).

[17] Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions - Building Trust in Human Centric Artificial Intelligence (COM(2019)16. https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=58496 (last access 26-04-2024).

[18] See, The Assessment List for Trustworthy Artificial Intelligence (ALTAI) for Self-assessment. https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=68342 (last access 14-04-24).

[19] White Paper on Artificial Intelligence-A European approach to excellence and trust. https://commission.europa.eu/document/download/d2ec4039-c5be-423a-81ef-b9e44e79825b_en?filename=commission-white-paper-artificial-intelligence-feb2020_en.pdf (last access 23-04-24).

[20] Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions A European Strategy for data COM/2020/66 final. https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52020DC0066 (last access 22-04-24).

[21] Inception impact assessment: Ethical and legal requirements on AI, https://ec.europa.eu/info/law/better-regulation/have-your-say/initiatives/12527-Artificial-intelligence-ethical-and-legal-requirements_en (last access 14-04-24).

[22] Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (artificial intelligence act) and amending certain union legislative acts COM/2021/206 final. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52021PC0206. (last access 20-04-24).

[23] Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions. Fostering a European approach to Artificial Intelligence, COM/2021/205 final. https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=COM%3A2021%3A205%3AFIN (last access 20-04-24).

focuses on adapting liability rules to the digital age and AI.[24] Thus, it was aimed at providing legal certainty and trust for investment in and societal acceptance of emerging technologies. The new Product Liability Directive attempted to ensure fair and predictable rules for businesses and victims, compensation for damages caused by unsafe products (including digital products and AI systems), and a level playing field between EU and non-EU manufacturers.[25] Previously, the EC released the Report on the safety and liability implications of AI, the Internet of Things, and robotics.[26]

In November 2021, the Slovenian presidency of the EU Council shared the first draft of the AIA, featuring changes to sections on social scoring, biometric recognition systems, and high-risk applications.[27] In December 2022, the Council adopted its common position on the AIA.[28] The aim is to ensure the safety of AI systems in the EU market while respecting fundamental rights and Union values. The proposal follows a risk-based approach and provides a uniform legal framework for AI to ensure legal certainty. It promotes investment and innovation in AI, enhances governance and enforcement of fundamental rights and safety, and facilitates the development of a single market for AI applications. This regulation aligns with other initiatives, including the Coordinated Plan on Artificial Intelligence, to accelerate investment in AI in Europe.

In June 2023, the EP adopted its negotiating position on the AIA, aiming to promote the uptake of human-centric and trustworthy AI while protecting health, safety, fundamental rights, and democracy. The rules ensure that AI developed and used in Europe aligns with EU rights and values, including human oversight, safety, privacy, transparency, non-discrimination, and social and environmental well-being. Key provisions include a full ban on AI for biometric surveillance (some exceptions are permitted based on compelling public interest: locating missing persons, protecting individuals from physical harm, or mitigating terrorist threats, etc.), emotion recognition, and predictive policing, as well as requirements for generative AI systems like ChatGPT to disclose their AI-generated content. Moreover, AI systems used to influence voters in elections are considered high-risk.

In August 2023, the Council of the EU published a document in preparation for the trialogue. In December 2023, the EP and the Council reached a political agreement on the AIA because of the interinstitutional talks (trilogue). In February 2024, representatives from member states unanimously voted to adopt the AIA on the basis of that political agreement. The committees of the EP (Internal Market and Consumer Protection 'IMCO' and Civil Liberties, Justice and Home Affairs 'LIBE') also voted to adopt the EU AIA. The final parliamentary vote took place in March 2024. The EP supported the regulation with 523 votes in favour, 46 against, and 49 abstentions, following the Ordinary Legislative Procedure. However, it is not yet in force. The AIA, in Article 113, outlines a progressive application schedule with a series of deadlines.[29]

Due to the extensive implications of AI, additional regulations related to data governance, cybersecurity, harmonisation of national liability rules for AI, intellectual property, and business competitiveness will be necessary. AI also affects sectors such as education. EU member states, with support from the Commission, should expedite efforts to retrain individuals and reform education systems to address evolving employment demands. By promoting skills such as creativity, communication, and critical thinking, citizens will be better equipped to navigate a changing job landscape. Funding should be directed towards universities and research labs to cultivate AI talent [11,17]. Then, to ensure a coherent regulatory landscape, it will achieve horizontal alignment with existing European legislation.

Finally, the EU has put together a comprehensive and relatively systematic AI regulation that provides for governance mechanisms to be able to implement the regulation effectively. By being the first to regulate, the EU can set the de facto global standard, shaping how the world does business and develops technology ("Brussels effect"). The EU's standards and domestic legislation have an extraterritorial impact. Thus, the AIA can set a global standard for AI regulation in other states, promoting a European approach to technology regulation on the world stage. Due to the global reach of AI that transcends state borders, its regulation demands regional

[24] Proposal for a Directive of the European Parliament and of the Council on adapting non-contractual civil liability rules to artificial intelligence (AI Liability Directive) COM/2022/496 final. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52022PC0496 (last access 22-04-24).

[25] Product Liability Directive - Adapting liability rules to the digital age, circular economy and global value chains. https://ec.europa.eu/info/law/better-regulation/have-your-say/initiatives/12979-Product-Liability-Directive-Adapting-liability-rules-to-the-digital-age-circular-economy-and-global-value-chains_en (last access 20-04-24)

[26] Report from the Commission to the European Parliament, the Council and the European Economic and Social Committee Report on the safety and liability implications of Artificial Intelligence, the Internet of Things and robotics, COM/2020/64 final. https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52020DC0064 (last access 20-04-24)

[27] Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts-Presidency compromise text. https://data.consilium.europa.eu/doc/document/ST-14278-2021-INIT/en/pdf (last access 20-04-24)

[28] Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts-General approach. https://data.consilium.europa.eu/doc/document/ST-14954-2022-INIT/en/pdf (last access 20-04-24).

[29] As indicated in a press release from the EP: "The regulation is still subject to a final lawyer-linguist check and is expected to be finally adopted before the end of the legislature (through the so-called corrigendum procedure). The law also needs to be formally endorsed by the Council. It will enter into force twenty days after its publication in the official Journal, and be fully applicable 24 months after its entry into force, except for: bans on prohibited practises, which will apply six months after the entry into force date; codes of practice (nine months after entry into force); general-purpose AI rules including governance (12 months after entry into force); and obligations for high-risk systems (36 months)". See, Artificial Intelligence Act: MEPs adopt landmark law. https://www.europarl.europa.eu/news/en/press-room/20240308IPR19015/artificial-intelligence-act-meps-adopt-landmark-law . (last access 25-04-24).

and global solutions for greater legal security and protection of people's rights. Guidelines and cooperation at the international level are essential to foster a global convergence of ethical principles and practices.[30]

## 4. Building collaborative AI governance in the EU

Talking about AI governance, we mean "a system of rules, practices, processes, and technological tools that are employed to ensure an organization's use of AI technologies aligns with the organization's strategies, objectives, and values; fulfills legal requirements; and meets principles of ethical AI followed by the organization" [23]. In addition to the development of legal and ethical frameworks (ethical guidelines, best practices or codes of conduct, etc.), the harmonised and effective implementation of the AIA requires well-defined governance mechanisms.[31] These mechanisms should primarily facilitate collaboration among institutions, society, and other stakeholders, enabling them to achieve policy objectives in a dynamic and evolving environment without significant disruptions or harm to society [35]. On the other hand, AI governance must be sensitive and consider the motivations and incentives of institutions, actors and stakeholders. It should balance the benefits and risks of AI, leveraging knowledge and a capacity to evaluate progress and implement incentives and sanctions that support desirable outcomes (Carsten, 2021).

The AIA includes Chapter VII (articles 64 to 70 AIA), which specifically outlines the main governance mechanisms. It is divided into two sections; Section 1 is dedicated to Governance at the Union level, while Section 2 addresses the National competent authorities. The first element of this governance system is the development of expertise and capabilities at the EU level, which is essential for fostering understanding, contributing to the implementation and enforcement of the Regulation on AI, and implementing international rules and principles on AI, such as the G7 Code of Conduct and Guiding Principles for developers of advanced AI systems.[32] Firstly, article 64 establishes the European AI Office, which aims to develop expertise and capabilities in the field of AI within the EU. The EC is responsible for establishing and managing this Office and Member States must support the tasks assigned to the AI Office.

To ensure timely preparation for the implementation of the AIA (art. 64 AIA), in February 2024 the European AI Office was created within the Commission to oversee the enforcement and implementation of the AIA with member states.[33] Overall, its main aim is to support the enforcement of the AIA, particularly regarding general-purpose AI. It strives to ensure that AI technologies respect human dignity, rights, and trust, while fostering collaboration, innovation, and research among stakeholders. Moreover, it engages in international dialogue and cooperation to align global AI governance, aiming to position Europe as a leader in the ethical and sustainable development of AI technologies. The European AI Office has been established as part of the administrative structure of the Directorate-General for Communication Networks, Content and Technology, and subject to its annual management plan (Decision art. 1). It is notable that it is not an agency.

In its article 2, the EC Decision establishes that the Office shall perform the tasks outlined in article 3 to implement and enforce the AIA. According to this latter article, concerning its implementation, the Office performs specified tasks derived from the AIA. Particularly, "1. The Office shall perform the following tasks stemming from the forthcoming Regulation:

(a) developing tools, methodologies and benchmarks for evaluating capabilities of general-purpose AI models, in particular for very large general purpose AI models with systemic risks;
(b) monitoring the implementation and application of rules on general-purpose AI models and systems, in particular where the model and the system are developed by the same provider;
(c) monitoring the emergence of unforeseen risks stemming from general-purpose AI models, including by responding to alerts from the scientific panel;
(d) investigating possible infringements of rules on general-purpose AI models and systems, including by collecting complaints and alerts, assisting in the preparation of decisions of the Commission and conducting evaluations pursuant to the forthcoming Regulation;
(e) ensuring that when an AI system falls within the scope of relevant Union legislation for which the Commission has powers of supervision and enforcement, such as under Regulation (EU) 2022/2065 of the EP and of the Council (7) or under Regulation (EU) 2022/1925 of the EP and of the Council (8), the supervision and enforcement of that legislation is fully coordinated with the supervision and enforcement of the forthcoming Regulation;

[30] In May 2023, G7 countries agreed to prioritize collaboration for inclusive AI governance. Governments emphasized the importance of adopting forward-looking, risk-based approaches to trustworthy AI, in line with shared democratic values. See G7 Hiroshima Leaders' Communiqué, https://www.whitehouse.gov/briefing-room/statements-releases/2023/05/20/g7-hiroshima-leaders-communique/

[31] On this issue, the European Commission opposed the creation of a new institution. Meanwhile, the European Parliament advocated for the establishment of a European Agency for Artificial Intelligence. See, Stahl, B. C. [33] Artificial Intelligence for a Better Future An Ecosystem Perspective on the Ethics of AI and Emerging Digital Technologies, Springer, https://doi.org/10.1007/978-3-030-69978-9

[32] Hiroshima Process International Code of Conduct for Organizations Developing Advanced AI Systems, 30 October 2023. See, https://ec.europa.eu/newsroom/dae/redirection/document/99641 (last access 25-04-24).

[33] Commission Decision establishing the European Artificial Intelligence Office, C(2024) 390 final. https://eur-lex.europa.eu/eli/C/2024/1459/oj (last access 20-04-24). On 21 February 2024, the European Artificial Intelligence Office was established in the Commission, reporting to the Directorate-General for Communication Networks, Content and Technology Directorate-General for Communications Networks, Content and Technology (DG CNECT). It was into force on 21 February 2024.

(f) supporting the implementation of rules on prohibited AI practices and high-risk AI systems in coordination with relevant bodies responsible under sectoral legislation, including facilitating information exchange and collaboration between national authorities, collecting notifications and establishing information platforms and databases, in particular when a general-purpose AI model or system is integrated into a high-risk AI system.

2. Moreover, in order to contribute to the effective implementation of the forthcoming Regulation, the Office shall be tasked with:

(a) assisting the Commission in the preparation of relevant Commission Decisions, and of implementing and delegating acts;
(b) facilitating the uniform application of the forthcoming Regulation; […]"
(c) coordinating the establishment of an effective governance system, including by preparing the set-up of advisory bodies at Union level, as well as monitoring the establishment of relevant national authorities and other bodies at national level;
(d) providing the Secretariat for the AI Board and its subgroups and providing administrative support to the advisory forum and the scientific panel, where applicable, including by providing the administrative set-up, organising meetings and preparing relevant documents; […]".

Notably, the AI Office will execute its duties through a collaborative approach, engaging with stakeholders, providers of AI models, relevant experts, and social entities. It also contemplates cooperation with other relevant bodies and departments to promote the societal and industrial benefits of AI technologies; especially, with the European Centre for Algorithmic Transparency. Along these lines, it also refers to inter-institutional cooperation (other bodies, offices, and agencies of the EU) and international cooperation.

Other governance mechanisms are provided in Chapter VII: Governance of the AI Regulation, which establishes the European Artificial Intelligence Board (art. 65 AIA). It is composed of one representative per Member State, with the European Data Protection Supervisor participating as an observer. The AI Office also attends the Board's meetings without voting rights, and other national and Union authorities or experts may be invited on a case-by-case basis. States' representatives are designated for three years, renewable once, must have relevant competences and powers, and act as single contact points for the Board and stakeholders. They must be empowered to facilitate consistency and coordination between national competent authorities in their Member State in order to implement the AIA. The Board selects the chair (it is chaired by one of the Member State representatives) and establishes two standing sub-groups for market surveillance and may create additional sub-groups as needed. It operates to ensure objectivity and impartiality. The AI Office provides secretariat support.

Article 66 AIA indicates that "the Board shall advise and assist the Commission and the Member States in order to facilitate the consistent and effective application of this Regulation". For example, the Board can facilitate coordination among national competent authorities, collect and share technical and regulatory expertise and best practices among Member States, provide advice on implementing the Regulation, particularly regarding enforcement of rules on general-purpose AI models, etc.

According to article 67, the Board will be assisted by an advisory forum to provide technical expertise and advice to both the Board and the Commission, supporting their tasks under the AIA. The advisory forum's membership will consist of a balanced selection of stakeholders, including representatives from industry, start-ups, SMEs, civil society, and academia. Membership will be balanced in terms of commercial and non-commercial interests, with special consideration given to SMEs and other commercial entities. Finally, article 68 outlines the establishment of a scientific panel of independent experts, intended to support enforcement activities under the AIA. The "scientific panel" will consist of experts selected by the Commission based on their up-to-date scientific or technical expertise in the field of AI. These experts shall carry out their tasks with impartiality and objectivity, ensuring the confidentiality of information and data obtained during their activities.

Lastly, article 70 AIA (Section 2: National competent authorities) refers to the designation of national competent authorities (and single point of contact). Particularly, it indicates that "each Member State shall establish or designate as national competent authorities at least one notifying authority and at least one market surveillance authority for the purposes of this Regulation [AIA]". Moreover, these will act independently, impartially and without bias to preserve the objectivity of their activities and guarantee application of the AIA. The competent national authorities must also respect confidentiality obligations. It also determines that Member States must ensure that national competent authorities are provided with adequate financial and human resources to fulfil their tasks arising from the AIA. Specifically, those authorities should have a sufficient number of personnel with expertise in AI technologies, data and data computing, personal data protection, cybersecurity, fundamental rights, health and safety risks, and knowledge of existing standards and legal requirements. The national competent authorities may also provide guidance and advice on the implementation of this Regulation taking into account the guidance and advice of the Board and the Commission, as appropriate.

In terms of governance, national competent authorities will oversee the implementation of the AIA rules at the national level, while the European AI Office will ensure coordination at the European level. Alongside the national authorities, the AI Office will be the first body globally to enforce binding rules on AI and is therefore expected to become an international reference point. This position will be strengthened due to its connection with the European Artificial Intelligence Board, advisory forum and the scientific panel that help incorporate the socio-economic and technical dimension (industry, scientific community, civil society and other stakeholders). Obviously, this governance mechanism must be capable of assuming the tasks entrusted to it and, above all, guarantee the implementation and compliance of the regulation. The AI Office was created to be the centre of AI expertise and the basis of a single European system of AI governance. In short, in terms of governance, a network of European and state authorities is established based on new specialized organizations (and on sectoral networks of regulatory authorities). Collaboration between national authorities among themselves and with the European Commission is a central element.

Ensuring consistent implementation of the AI regulations demands a single and centralized governance approach. To achieve this,

the AI Office, AI Board, and advisory forum should act as central facilitators. This necessitates fostering continuous engagement with industry, experts, and stakeholders through effective communication channels. Furthermore, collaboration with national authorities responsible for AI governance is crucial for ensuring coherence across different jurisdictions within the EU. By fostering a collaborative and centralized approach, the EU can strive to achieve effective and consistent AI governance.

## 5. Conclusions

As we indicated, the main objective of this text is to critically examine the governance system established by the AIA. However, a significant limitation is the lack of practical implementation. For this reason, we have limited our attention to the contents of the current AIA version.

As we have observed, AI presents a range of challenges and can have adverse effects. Although some challenges may be considered independently, many are highly interconnected, giving rise to social, ethical, and legal issues simultaneously. By approving the AIA, the EU takes a view emphasizing the role of governments in ensuring value to stakeholders. It assumes that government agencies should use various functions to create a positive impact for all those involved in the development and use of AI. So, as Birkstedt et al. [8] points out, AI governance "is about how humans can best advance AI development, which comprises three main elements: technical landscape (limits and scope of AI), ideal governance (potential pathways for facilitating stakeholder cooperation) and AI politics (political dynamics affecting stakeholders)".

The EU's governance mechanism responds to a holistic or comprehensive view based on these strong interconnections. As Li pointed out "interdisciplinary collaboration and stakeholder engagement are crucial for addressing the multifaceted challenges and ethical considerations associated with AI development and deployment. Collaboration among computer scientists, ethicists, policymakers, and industry stakeholders facilitates the integration of diverse perspectives, expertise, and values in the design, implementation, and governance of AI technologies" [22].

In the same way, Sonko et al. [32] indicates that a comprehensive approach to AI development involves integrating technical, ethical, and societal considerations throughout the research and deployment process. This requires recognizing the interconnectedness of technical advancements, ethical principles, and societal impacts, and adopting an interdisciplinary perspective that accounts for diverse perspectives, values, and priorities. By taking a holistic approach, we can address the complex challenges and ethical dilemmas associated with AI in a comprehensive and systematic manner, ensuring that AI technologies are developed and deployed in ways that benefit humanity while minimizing risks and negative consequences.

The EU has established governance mechanisms to address the challenges associated with AI and prevent personal and social harm, while building trust and social acceptance of AI. These mechanisms are based on coordination between European and state public authorities and on the collaboration of public and private actors. These elements are key factors to successfully implement regulation in the complex field of AI. Furthermore, like other governance mechanisms, EU instruments must be able to react quickly and flexibly to changes. Policymakers will evaluate technological advances in AI and be able to respond quickly with appropriate measures. As Carsten (2021) wrote "governance needs to provide spaces for actors to learn and develop understanding of technologies, their use and their evaluation. It must be based on and foster communication between stakeholders. It should also allow for the acknowledgement of mistakes and have the ability to reverse or change course where initial assumptions prove to be wrong or where new insights or consensus emerge".

As we have seen, the governance mechanisms provided for in the AIA respond to a collaborative logic by bringing in representatives from state administrations or experts from private organizations. In this way, these representatives and experts, in addition to developing regulations and standards, provide updated information and knowledge on current and future AI issues and make new proposals. Communication flows between public organizations and private entities are also favoured, which is of utmost importance in a field as dynamic as AI.

However, there are academic voices critical of these collaborative governance mechanisms. For instance, Lévesque [21] indicates that "while collaborative AI governance aspires to actively involve interested and impacted stakeholders upstream, the concentration of power in a few industry players veers into self-governance. […] AI regulation purporting to embody collaborative governance actually facilitate [*sic*] self-governance by regulated entities […]. [It] takes issue with the disconnect between collaborative governance's optically palatable messaging that states smartly include everyone in AI regulation, and the actual mechanics of legislation effectively insulating AI companies from public oversight.". According to the same author, this assertion rests on two key points. Firstly, established industry players hold undue sway over supposedly participatory processes. Secondly, lacunae within the proposed legislation incentivise regulated entities to circumvent the most onerous tier of obligations [21].

In this same sense, Birkstadt *et al.* [8] point out a power imbalance between those who set the rules for AI systems (regulations, norms, and client demands) and the developers who implement them. This imbalance can lead to biased application of AI governance principles, potentially favouring private organizations, especially in developed economies. Besides the unequal allocation of technical expertise and power, they also point to a hierarchy among the different stakeholders of AIG [AI Governance] in terms of their degrees of influence, specific responsibilities and accountability levels.

In conclusion, a critical evaluation of the evolving dynamics and implementation patterns within the collaborative AI governance framework is essential. This ongoing scrutiny will be crucial in striking a balance between safeguarding European citizens' rights and fostering a competitive and innovative AI industry in Europe.

## CRediT authorship contribution statement

**Celso Cancela-Outeda:** Writing – review & editing, Writing – original draft, Conceptualization.

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Data availability

No data was used for the research described in the article.

## References


[1] W. Abbaoui, S. Retal, B. El Bhiri, N. Kharmoum, S. Ziti, Towards revolutionizing precision healthcare: a systematic literature review of artificial intelligence methods in precision medicine, Inform. Med. Unlocked. 46 (2024) 1–19, https://doi.org/10.1016/j.imu.2024.101475.

[2] P.S. Aithal, Super-intelligent machines-analysis of developmental challenges and predicted negative consequences, Int. j. appl. eng. manag. lett. (IJAEML) 7 (3) (2023) 109–141.

[3] M. Alier, F.J. García-Peñalvo, J.D. Camba, Generative artificial intelligence in education: from deceptive to disruptive, Int. j. interact. multimed. artif. intell. 8 (2024) 5–14. *Special issue on Generative Artificial Intelligence in Education5*.

[4] T. Babina, A. Fedyk, A. He, J. Hodson, Artificial intelligence, firm growth, and product innovation, J. financ. econ. 151 (2024), https://doi.org/10.1016/j.jfineco.2023.103745.

[5] Baily, M.T. and Kane, A. (2 May 2024), How will AI affect productivity? *Brookings*. https://www.brookings.edu/articles/how-will-ai-affect-productivity/(last access 10 May 2024).

[6] Z. Bednarz, La inteligencia artificial responsable para el sector financiero: consideraciones legales, Revista de Derecho del Mercado de Valores 29 (2021) 8.

[7] C. Bianchi, G. Nasi, W.C. Rivenbark, Implementing collaborative governance: models, experiences, and challenges, Public Management Rev. 23 (11) (2021) 1581–1589, https://doi.org/10.1080/14719037.2021.1878777.

[8] T. Birkstedt, M. Minkkinen, A. Tandon, M. Mäntymäki, AI governance: themes, knowledge gaps and future agendas, Internet Res. 33 (7) (2023) 133–167.

[9] C. Cancela-Outeda, La Inteligencia artificial: Una Lectura Desde La política. Inteligencia artificial: De La Discrepancia Regional a Las Reglas universales. Integración de Percepciones políticas, Económicas y Legales, Thomson Reuters Aranzadi, 2020, pp. 319–344.

[10] C. Cancela-Outeda, B. González-Cacheda, Artificial Intelligence: a Reading from European Politics, in: D. Ramiro Troitiño, T. Kerikmäe, O. Hamuľák (Eds.), Digital Development of the European Union, Springer, Cham, 2023, https://doi.org/10.1007/978-3-031-27312-4_23.

[11] G. Carriço, The EU and artificial intelligence: a human-centred perspective, European View 17 (1) (2018) 29–36, https://doi.org/10.1177/1781685818764821.

[12] J. Collenette, K. Atkinson, T. Bench-Capon, Explainable AI tools for legal reasoning about cases: a study on the european court of human rights, Artif. Intell. 317 (2023) 103861, https://doi.org/10.1016/j.artint.2023.103861.

[13] M.Á. Escotet, The optimistic future of artificial intelligence in higher education, Prospects. (2023), https://doi.org/10.1007/s11125-023-09642-z.

[14] J. Franganillo, La inteligencia artificial generativa y su impacto en la creación de contenidos mediáticos, Methaodos rev. cienc. soc. 11 (2) (2023) m231102a10, https://doi.org/10.17502/mrcs.v11i2.710.

[15] GAO. U. S. Government Accountability Office (6 September 2023). Artificial Intelligence's Use and Rapid Growth Highlight Its Possibilities and Perils. *WatchBlog*. https://www.gao.gov/blog/artificial-intelligences-use-and-rapid-growth-highlight-its-possibilities-and-perils (last access 10 May 2024).

[16] B. Girish, M. Rajvi, Review on applications of artificial intelligence in agriculture, EPRA Int. J. Agric. Econ. Res. (ARER) 12 (1) (2024) 12–17, https://doi.org/10.36713/epra0813.

[17] L. Hamilton, D. Elliott, A. Quick, S. Smith, V. Choplin, Exploring the Use of AI in qualitative analysis: a comparative study of guaranteed income data, Int. J. Qual. Methods 22 (2023), https://doi.org/10.1177/16094069231201504.

[18] Hendrycks, D., Mazeika, M., & Woodside, T. (2023). An overview of catastrophic ai risks. https://arxiv.org/abs/2306.12001.

[19] J. Hertzberg, B. Kisliuk, J.C. Krause, AI in current and future agriculture, Künstl Intell 37 (2023) 113–115, https://doi.org/10.1007/s13218-024-00838-9.

[20] L. Inglada, L. Corral, P. Miramontes, Ethics and artificial intelligence, Revista Clínica Española 224 (3) (2024) 178–186, https://doi.org/10.1016/j.rce.2024.01.007.

[21] Lévesque, M. (22 February 2024). Collaborative AI governance: minding the gap between theory and practice. available at SSRN: https://ssrn.com/abstract=4736131 or https://doi.org/10.2139/ssrn.4736131.

[22] N. Li, Ethical considerations in artificial intelligence: a comprehensive disccusion from the perspective of computer vision, SHS Web Conf. 179 (1) (2023), https://doi.org/10.1051/shsconf/202317904024.

[23] M. Mäntymäki, M. Minkkinen, T. Birkstedt, M. Viljanen, Defining organizational AI governance, AI. Ethics 2 (4) (2022) 603–609, https://doi.org/10.1007/s43681-022-00143-x.

[24] P. Mikalef, K. Conboy, J.E. Lundström, A. Popovič, Thinking responsibly about responsible AI and 'the dark side' of AI, Eur. J. Inf. Syst. 31 (3) (2022) 257–268, https://doi.org/10.1080/0960085X.2022.2026621.

[25] A. Nikitas, K. Michalakopoulou, E.T. Njoya, D. Karampatzakis, Artificial intelligence, transport and the smart city: definitions and dimensions of a new mobility era, Sustainability. 12 (2020) 2789, https://doi.org/10.3390/su12072789.

[26] C. Paz-y-Miño, La inteligencia artificial en medicina general y en genómica, Metro Ciencia 31 (2) (2023) 81–86, https://doi.org/10.47464/MetroCiencia/vol31/2/2023/81-86.

[27] D. Ramiro Troitiño, The European Commission, the Council, and the European Parliament: differentiated Theoretical Frame for the Digital Revolution. book: Digital Development of the European Union, An Interdisciplinary Perspective Publisher: Springer, Cham, 2023, https://doi.org/10.1007/978-3-031-27312-4_22.

[28] T. Roy, The rapid rise of AI: advancements, concerns, and future outlook, All Tech Magazine (2 October 2023). https://alltechmagazine.com/the-rapid-rise-of-ai-advancements-concerns-and-future-outlook/ (last access 25-04-2024).

[29] T. Ryan-Mosley, Vuelta al mundo por las regulaciones de la IA en 2024, MIT Tech. Rev. (23 January 2024). https://www.technologyreview.es/s/16069/vuelta-al-mundo-por-las-regulaciones-de-la-ia-en-2024 (last access 30-04-24).

[30] y M. Sánchez, E. Carbajal, La inteligencia artificial generativa y la educación universitaria: ¿Salió el genio de la lámpara?, Perfiles Educativos 45 (2023) 70–86, https://doi.org/10.22201/iisue.24486167e.2023.Especial.61692 (Especial).

[31] Singh, T. (27 June 2023). *AI Act: how to balance productivity* vs. *safety?* [Post de LinkedIn]. LinkedIn. https://www.linkedin.com/posts/tarrysingh_ai-act-how-to-balance-productivity-vs-safety-123456789.

[32] S. Sonko, A. Adewusi, O. Obi, S. Onwusinkwue, A. Atadoga, A critical review towards artificial general intelligence: challenges, ethical considerations, and the path forward, World J. Adv. Res. Rev. 21 (03) (2024) 1262–1268, https://doi.org/10.30574/wjarr.2024.21.3.0817.

[33] B.C. Stahl, Artificial Intelligence for a Better Future. An Ecosystem Perspective on the Ethics of AI and Emerging Digital Technologies, Springer, 2021, https://doi.org/10.1007/978-3-030-69978-9.
[34] M. Veale, K. Matus, R. Gorwa, AI and global governance: modalities, rationales, tensions, Annu Rev. Law Soc. Sci. 19 (2023) 255–275, https://doi.org/10.1146/annurev-lawsocsci-020223-040749.
[35] B.W. Wirtz, J.C. Weyerer, B.J. Sturm, The dark sides of artificial intelligence: an integrated AI governance framework for public administration, Int. J. Public Adm. 43 (9) (2020) 818–829, https://doi.org/10.1080/01900692.2020.1749851.